

拟南芥未知功能基因 *At4g22890* 编码蛋白的叶绿体定位*

崔永兰¹, 王鹏程^{1,2}, 杨仲南^{1**}

(1 上海师范大学生命与环境科学学院, 上海 200234; 2 烟台大学体育学院, 山东 烟台 264005)

摘要: 蛋白质的亚细胞定位信息对于深入了解该蛋白质的功能具有重要意义。本文对一个预测的拟南芥叶绿体未知功能基因 *At4g22890* 编码蛋白进行了叶绿体定位研究。我们克隆了该基因 5 端长 208 bp 的 DNA 片段, 与绿色荧光蛋白 (GFP) 基因构建重组表达载体 pMON530-cTP-GFP, 经农杆菌介导转化拟南芥。转基因植株经激光共聚焦显微镜观察, GFP 荧光仅在叶绿体中观察到, 表明所克隆的 DNA 序列编码的多肽能够将 *At4g22890* 编码蛋白质引导进入叶绿体, 由此推测该蛋白质为叶绿体蛋白质。

关键词: 拟南芥; 转运肽; 融合蛋白; 叶绿体定位

中图分类号: Q 943

文献标识码: A

文章编号: 0253-2700 (2007) 06- 659- 06

Chloroplast Subcellular Localization of the Expressed Protein
Encoded by *At4g22890* in *Arabidopsis thaliana* (Cruciferae)

CUI Yong-Lan¹, WANG Peng-Cheng^{1,2}, YANG Zhong-Nan^{1**}

(1 College of Life and Envionment Science, Shanghai Normal University, Shanghai 200234, China;

2 College of Physical Education, Yantai University, Yantai 264005, China)

Abstract: The information of protein subcellular localization is important to our understanding of a protein function . The protein encoded by *At4g22890* was predicted to be an *Arabidopsis* chloroplast protein . The 5 208 bp fragment of this gene was cloned and fused with GFP to construct a binary vector pMON530-cTP-GFP for genetic transformation . On confocal laser-scanning microscopy, green fluorescent signals were localized in chloroplasts in transgenic *Arabidopsis* plants, suggesting that the 208 bp DNA fragment encodes a peptide targeting to chloroplast . The results suggest that *At4g22890* encodes a chloroplast protein .

Key words: *Arabidopsis thaliana*; Transit peptide; Fusion protein; Chloroplast localization

叶绿体是高等植物及藻类进行光合作用的场所, 是一个非常复杂的代谢单位, 除了光合作用的全部过程 (光的吸收、电荷分离、电子传递、光合磷酸化, CO₂ 的固定与还原等) 之外, 同时还参与氨基酸、脂肪酸、植物生长物质、核苷酸、维生素及次生代谢产物的合成 (Motohashi 等, 2001)。这些重要生化过程的进行依赖于各种各样的蛋白质, 这些蛋白或是色素 - 蛋白质复合体中不可缺少的成员, 或是电子载体, 或是蛋

白质转运体, 或是酶。叶绿体中的蛋白质只有 60~80 个为叶绿体基因组编码, 其余均为核基因组编码。

核基因编码的叶绿体蛋白的 N 端往往具有叶绿体转运肽序列特征, 通过对基因编码蛋白质的 N 端序列进行分析, 可以预测定位于叶绿体中的蛋白质。生物信息学分析表明, 模式植物拟南芥中有 4255 个核基因编码的叶绿体蛋白质 (Friso 等, 2004)。其中一部分通过蛋白质组学已得到分

* 基金项目: 科技部重大基础研究前期研究专项 (2003CCA01100) 及上海市教委项目 (06DZ016) 资助

** 通讯作者: Author for correspondence . E-mail: znyang@shnu.edu.cn

收稿日期: 2007-03-14, 2007-08-07 接受发表

作者简介: 崔永兰 (1974-) 女, 博士, 主要从事植物叶绿体发育方面的研究。

离。Kieselbach 等 (2000) 确定了 2 个类囊体蛋白质; Peltier 等 (2002) 分离出 81 个类囊体内腔和外周蛋白; 同年 Schubert 等 (2002) 又分离出 36 个类囊体内腔蛋白质; Ferro 等 (2003) 利用高效液相色谱技术分离出 112 个叶绿体被膜蛋白质; Froehlich 等 (2003) 利用双相电泳分离了 392 个叶绿体被膜蛋白; Friso 等 (2004) 采用多种方法分离出 154 个类囊体膜蛋白; Kleffmann 等 (2004) 采用串联质谱技术分离出 690 个叶绿体蛋白。

不管是生物信息学预测还是大规模蛋白质组学获得的蛋白质叶绿体定位信息, 都存在一定程度的错误率。蛋白质叶绿体定位的最终确定需要实验的证实。关于蛋白质的叶绿体定位已有成功报道, 如 Motohashi 等 (2001) 以 *APG2* 基因的 N 端转运肽序列与 GFP 构建融合蛋白研究该基因的亚细胞定位, 王鹏程等 (2006) 将预测的叶绿体未知功能蛋白 At5g48790 N 端转运肽与 GFP 融合, 成功地将 At5g48790 编码蛋白定位于叶绿体中。为了寻找与叶绿体发育及光合作用相关的基因, 我们利用反向遗传学的方法研究生物信息学预测的核基因组编码的未知功能基因 At4g22890。作为基因功能研究的一部分, 本文对该蛋白的亚细胞定位进行试验验证, 为进一步深入研究编码该蛋白的基因功能奠定基础。

1 材料与方法

1.1 材料

1.1.1 植物材料 野生型拟南芥 (*Arabidopsis thaliana*) Col-0 生态型, 由本实验室种植。培养室条件为: 22℃, 16 h 光照 8 h 黑暗, 光通量密度 110 μmol·m⁻²·s⁻¹, 湿度约为 30%。

1.1.2 菌株、质粒与 cDNA 大肠杆菌 (*Escherichia coli*) DH5 菌株由本实验室提供, 农杆菌 ASE 菌株、含 GFP 基因的质粒 pEGFP 与植物双元表达载体 pMON530 为中国科学院上海植物生理生态研究所黄海研究员惠赠, 基因 At4g22890 的全长 cDNA 克隆购自日本生物资源中心 RIKEN BRC (<http://www.brc.riken.jp/lab/epd/Eng/>)。

1.1.3 试剂 各种生化试剂和酶分别购自上海生工、大连宝生物和 Sigma 等公司。

1.1.4 引物合成 包含基因 At4g22890 N 端转运肽片段的 5 端引物序列为: ATG GGT AGC A AG ATG TTG; 3 端引物序列为: GTC CAT GGC TAC TGG ACC TGA TTG CT (下划线表示引入 NcoI 位点); GFP 内部引物序列 5 端

为 GTA AAC GGC CAC AAG TTC AGC G; 3 端为 GAT GCC GTT CTT CTG CTT GTC G。

1.2 方法

1.2.1 质粒提取、DNA 片段克隆与重组质粒鉴定 参照分子克隆实验指南 (黄培堂等译, 2002)。

1.2.2 DNA 片段回收 采用 BioDev 公司的胶回收试剂盒, 按其说明书进行操作。

1.2.3 农杆菌介导拟南芥的转化 采用浸花法 (Clough and Bent, 1998) 进行转化。

1.2.4 转基因 T₀ 代种子筛选 种子用 70% 乙醇 + 0.01% Triton X-100 混合液浸泡 10 min, 无菌水冲洗 4 次, 用 Top agar (0.1% 琼脂水溶液) 均匀涂布在具卡那霉素抗性 (50 μg·mL⁻¹) 的 PNS 固体培养基 (成分: 5 mmol/L KNO₃, 2.0 mmol/L Ca (NO₃)₂, 2.0 mmol/L MgSO₄, 50 mmol/L Fe-EDTA, 2.5 mmol/L KH₂ PO₄ (adjusted to pH 6.5), 70 mmol/L H₃ BO₃, 0.5 mmol/L CuSO₄, 0.2 mmol/L Na₂ MoO₄, 0.01 mmol/L CoCl₂, 10 mmol/L NaCl, 1 mmol/L ZnSO₄, 14 mmol/L MnCl₂, 0.7% Agar) 上, 4℃ 春化 2~3 d 后, 移入培养室培养。

1.2.5 转基因植株的 PCR 检测 以筛选出来的 T₁ 代植株的总 DNA 为模板, 采用 GFP 引物进行 PCR 扩增。

1.2.6 转基因植株的激光扫描共聚焦显微镜观察 用激光扫描共聚焦显微镜 (LSM 5 Pascal, Zeiss) 观察叶肉细胞中 GFP 的表达, 采用 40 倍油镜进行观察, 激发光波长为 488 nm, 带通 BP 为 505~530 nm, 长通 LP 为 560 nm。

2 结果与分析

2.1 叶绿体定位双元表达载体的构建

利用 *Kpn* 和 *EcoR* 限制性内切酶对 pEGFP 进行酶切, 获得包含 GFP 基因的 750 bp 片段, 将其插入到用相同限制性内切酶酶切的质粒 pMON530 中, 构建用于叶绿体定位的双元载体 pMON530-GFP (图 1: A, B)。

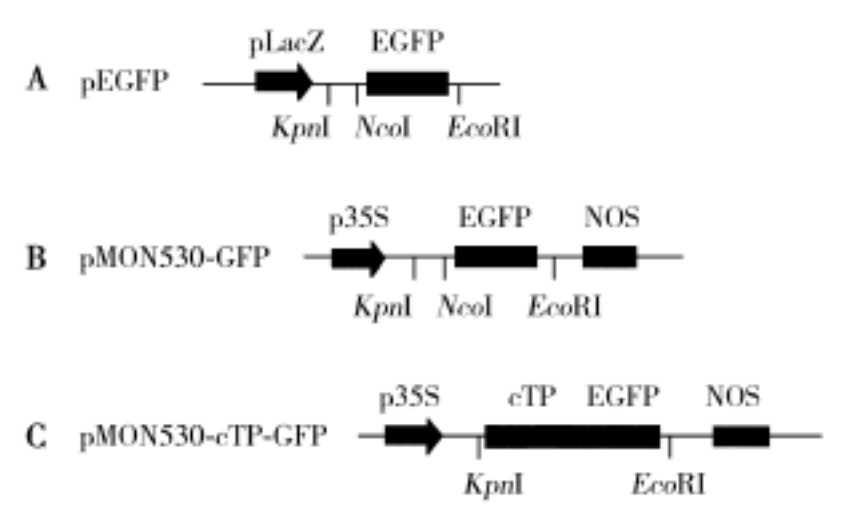


图 1 叶绿体定位表达载体 pMON530-cTP-GFP 的构建
Fig. 1 Construction of the expression vector pMON530-cTP-GFP

2.2 *At4g22890* 基因编码蛋白叶绿体定位表达载体的构建

我们通过 TargetP (<http://www.cbs.dtu.dk/services/TargetP>) 预测, 发现未知功能基因 *At4g22890* 编码蛋白质的 N - 端 60aa 具有转运肽结构。以 *At4g22890* 基因全长 cDNA (Seki 等, 1998; 2002) 为模板, PCR 扩增一段包含转运肽的 DNA 序列, 长度为 208 bp, 将其命名为 cTP。cTP 与 pMD18-T Vector 连接, 构建 T-A 克隆并经 M13 引物鉴定 cTP 的插入方向, 将正向插入克隆命名为 cTP-T。cTP-T 质粒经过 *Kpn* 和 *Noc* 双酶切与用相同酶切过的载体 pMON530-GFP 连接, 构建成表达载体 pMON530-cTP-GFP (图 1: C), 该表达载体经 *Kpn* 与 *EcoR* 双酶切验证, 结果表明插入 DNA 片段大小正确。

2.3 转基因植株的获得

将质粒 pMON530-GFP 与 pMON530-cTP-GFP 分别导入农杆菌 ASE, 采用农杆菌介导法转化拟南芥 Col-0 生态型。T₀ 代种子萌发经卡那霉素抗性筛选后, 分别得到 T₁ 代植株 25 棵和 20 棵, 转化效率分别为 0.8% 和 0.6%。经移栽后, 转基因植株表型正常。

2.4 转基因植株的 PCR 检测

对筛选出来的 T₁ 代植株各取 20 株采用 GFP 内部引物进行分子检测。在转基因 pMON530-cTP-GFP 植株中, PCR 检测有 19 株为阳性, 假阳性比例为 5% (图 2 给出部分结果); 在转基因 pMON530-GFP 植株中, PCR 检测 20 株全部为阳性。

2.5 报告基因 GFP 表达

对 PCR 检测阳性的植株用激光共聚焦显微镜观察其叶片中 GFP 的表达情况, 结果显示在 488 nm 的激发光下, 野生型 Col-0 仅在大于 560 nm 处出现叶绿素自发荧光 (红色), 在 GFP 发射光波段 505 ~ 530 nm 没有任何信号 (图 3: A ~ C)。pMON530-GFP 转基因植株在 505 ~ 530 nm 波段, GFP 荧光 (绿色) 在细胞质中呈弥散状分布 (图 3: D ~ F)。而在 pMON530-cTP-GFP 转基因植株中, GFP 荧光信号与叶绿素自发荧光信号共定位于叶绿体中 (图 3: G ~ I)。该结果表明基因 *At4g22890* 编码蛋白的 N 端 60aa 能够将该蛋白质引导进入叶绿体, 该蛋白质为核基因编码的叶绿体蛋白质。

2.6 *At4g22890* 基因编码蛋白质的转膜结构分析

基因 *At4g22890* 编码一条 324 个氨基酸组成的多肽链, 分子量为 35.7 kDa, 等电点为 5.20 (http://www.expasy.ch/tools/pi_tool.html), 为酸性蛋白质。通过转膜螺旋预测分析显示在两个疏水区域处相应地存在两个穿膜结构, 而 1 ~ 201aa, 220 ~ 238aa, 262 ~ 324aa 三个肽段可能处于膜系统的外侧 (<http://www.cbs.dtu.dk/services/TMHMM>) (图 4), 而疏水性分析结果显示 *At4g22890* 编码蛋白的 202 ~ 219aa 及 239 ~ 261aa 处存在两个疏水性区域 (<http://www.expasy.ch/tools/protscale.html>) (图5)。类囊体膜蛋白都含有叶绿体转运肽, 具有高度疏水性特点, 而且通常为偏酸性蛋白质 (Ferro 等, 2002)。结合 GFP 融合蛋白荧光定位情况 (图 3: G ~ I), 基因 *At4g22890* 编码的蛋白质具有类囊体膜蛋白的特性, 推测该蛋白质可能定位于叶绿体类囊体膜上。

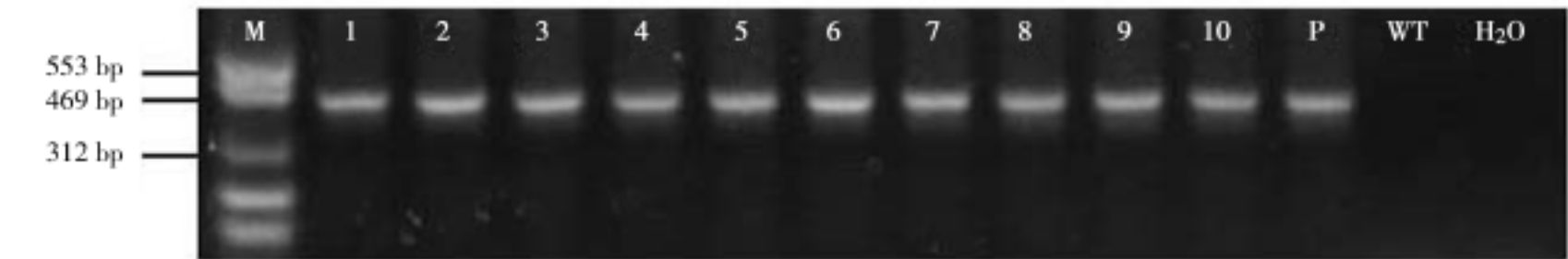


图 2 pMON530-cTP-GFP 转基因植株 PCR 检测

M . DNA 标准分子量 ; 1 ~ 10 . 转基因植株 ; P . pEGFP 质粒 (正对照);
WT . 野生型植株 (负对照); H₂O . 水 (负对照)

Fig . 2 PCR analysis of the pMON530-cTP-GFP transgenic plants

M . DNA Marker; lanes1 - 10 . Transgenic plants; P . Plasmid of pEGFP as positive control;
WT . Wild type plant as negative control; H₂O . H₂O as negative control

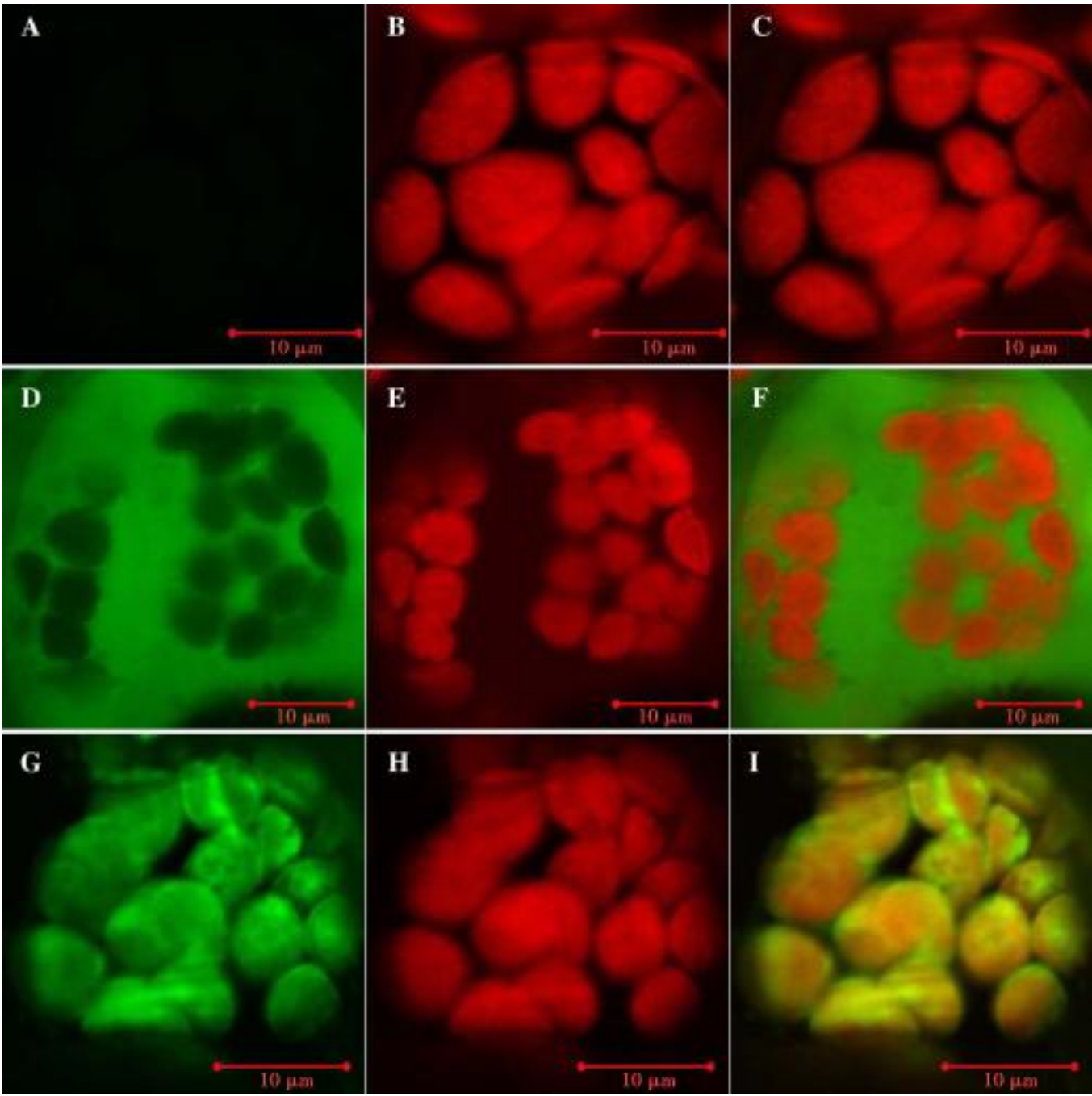


图 3 转基因拟南芥叶肉细胞中 GFP 荧光的观察

A ~ C . 野生型 Col-0 植株 GFP 荧光观察； D ~ F . 转基因 pMON530-GFP 植株 GFP 荧光观察； G ~ I 转基因 pMON530-cTP-GFP 植株 GFP 荧光观察； 其中， A、 D、 G 为 GFP 绿色荧光， B、 E、 H 为叶绿体自发荧光， C、 F、 I 分别为叠加后的结果； 注： A ~ I 图中标尺为 10 μm

Fig . 3 The fluorescence detection in the transgenic *Arabidopsis* cells

A-C . Wild type Col-0 fluorescent detection； D-F . Transgenic plants pMON-530 fluorescent detection； G-I . Transgenic plants pMON530-cTP-GFP fluorescent detection； A, D, G . GFP fluorescence； B, E, H . Chlorophyll autofluorescence； C, F, I . Overlapping . Note: scale bars , 10 μm

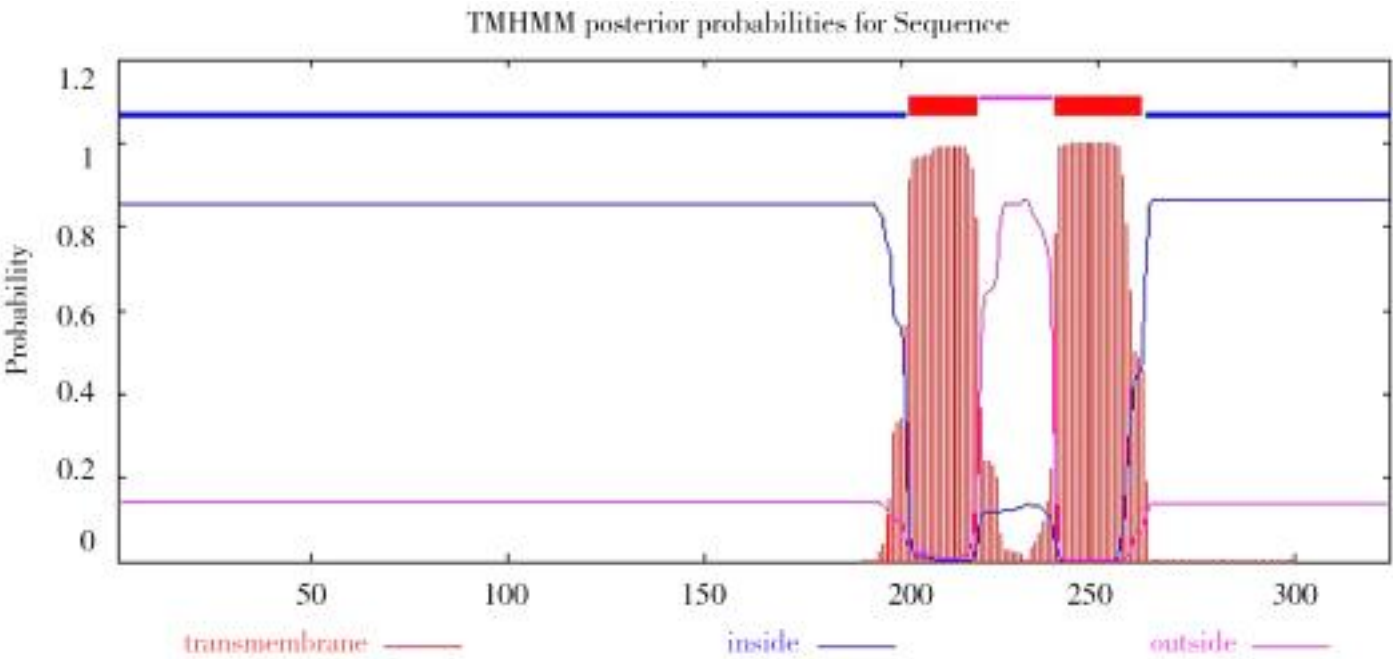
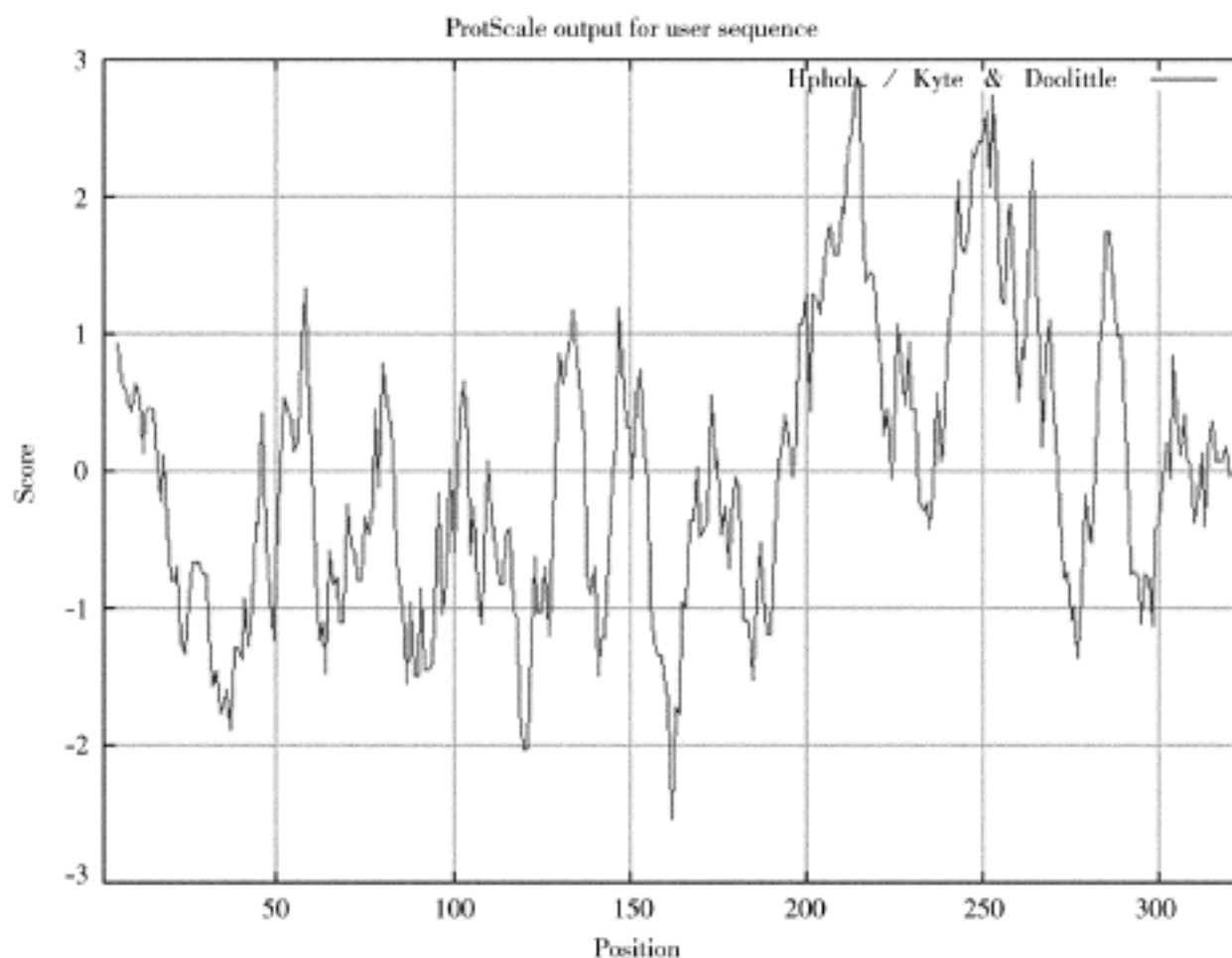


图 4 *At4g22890* 编码蛋白质的转膜螺旋结构预测

Fig . 4 Prediction of transmembrane helices in protein encoded by *At4g22890*

图 5 *At4g22890* 编码蛋白质的疏水性分析Fig. 5 Hydrophobic analysis of protein encoded by *At4g22890*

3 讨论

蛋白质的亚细胞定位对于阐明其功能具有重要作用。虽然通过生物信息学预测的蛋白质有 4 000 多种，但通过蛋白质组学分离的只有约 1 200 种，而真正通过实验验证的只有约 600 种。对于那些在叶绿体发育与光合作用中具有重要功能但丰度极低的蛋白质，通过蛋白质组学的方法进行分离则较为困难。同时由于亚细胞蛋白质分离过程中极易造成的污染使得通过蛋白质组学方法进行亚细胞定位的数据错误率较高。而采用体内表达目的蛋白质与荧光素的融合蛋白，则更为直观，而且能够进行活体检测，更能体现蛋白质在真实的生理状态下的定位情况。在 SUBA (<http://www.suba.bcs.uwa.edu.au>) 收集的亚细胞定位的数据中，其中在质体中通过荧光素定位的蛋白质有 142 个 (Heazlewood 等, 2006)。

通过转膜结构分析，我们推测 *At4g22890* 编码的蛋白质定位于拟南芥叶绿体类囊体膜上，而类囊体膜富集着很多与光合作用相关的蛋白质复合物或电子载体，*At4g22890* 编码蛋白质是否为其中的成员之一，究竟担负着怎样的生理功能，将来可通过购买该基因位点的 T-DNA 或转座子插入突变体或构建该位点基因敲除突变体来进一步阐明。

〔参 考 文 献〕

- 黄培堂等译 (Sambrook J, Russell DW, 2001), 2002. 分子克隆实验指南 (Molecular Cloning: A Laboratory Manual, 3rd ed.) [M]. 北京: 科学出版社
- Clough SJ, Bent AF, 1998. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* [J]. *Plant J*, **16** (6): 735—743
- Ferro M, Salvi D, Rolland HR *et al*. 2002. Integral membrane proteins of the chloroplast envelope: Identification and subcellular localization of new transporters [J]. *PNAS*, **99** (17): 11487—11492
- Ferro M, Salvi D, Brugiere S *et al*. 2003. Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana* [J]. *Mol Cell Proteomics*, **2** (5): 325—345
- Friso G, Giacomelli L, Ytterberg AJ *et al*. 2004. In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: New proteins, new functions, and a plastid proteome database [J]. *The Plant Cell*, **16** (2): 478—499
- Froehlich JE, Wilkerson CG, Ray WK *et al*. 2003. Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis [J]. *Proteome Res*, **2** (4): 413—425
- Heazlewood JL, Verboom RE, Tonti-Filippini J *et al*. 2006. SUBA: the *Arabidopsis* Subcellular Database [J]. *Nucleic Acids Res*, **00**: D1-D6
- Kieselbach T, Bystedt M, Hynds P *et al*. 2000. A peroxidase homologue and novel plastocyanin located by proteomics to the *Arabidopsis* chloroplast thylakoid lumen [J]. *FEBS Lett*, **480** (2-3): 271—276

- Kleffmann T, Russenberger D, Von Zychlinski A *et al* . 2004 . The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions [J] . *Curr Biol*, **14** (5): 354—362
- Motohashi R, Nagata N, Ito T *et al* . 2001 . An essential role of a TatC homologue of a Delta pH-dependent protein transporter in thylakoid membrane formation during chloroplast development in *Arabidopsis thaliana* [J] . *PNAS*, **98** (18): 10499—10504
- Peltier JB, Emanuelsson O, Kalume DE *et al* . 2002 . Central functions of the lumenal and peripheral thylakoid proteome of *Arabidopsis* determined by experimentation and genome-wide prediction [J] . *The Plant Cell*, **14** (1): 211—236
- Schubert M, Petersson UA, Haas BJ *et al* . 2002 . Proteome map of the chloroplast lumen of *Arabidopsis thaliana* [J] . *J Biol Chem*, **277** (10): 8354—8365
- Seki M, Carninci P, Nishiyama Y *et al* . 1998 . High-efficiency cloning of *Arabidopsis* full-length cDNA by biotinylated CAP trapper [J] . *Plant J*, **15** (5): 707—720
- Seki M, Narusaka M, Kamiya A *et al* . 2002 . Functional annotation of a Full-Length *Arabidopsis* cDNA collection [J] . *Science*, **296** (5565): 141—145
- Wang PC (王鹏程), Wang C (王晨), Mi HL (米华玲) *et al* . 2006 . Study of subcellular localization of the expressed protein in *Arabidopsis thaliana* [J] . *Chin Bull Bot* (植物学通报), **23** (3): 249—254

* * * * *

读者来信

对《水杉的未尽事宜》一文的几点看法

尊敬的《云南植物研究》主编及编辑同志：

你们好！

马金双先生在贵刊 2006 年第 5 期发表了《水杉的未尽事宜》一文，看到该文我很高兴。由于该文涉及几个我个人较为关心的问题，故欲借此机会谈点我的看法。不过，在入正题之前，先谈一下我对贵刊的态度。作为一个普通的学者，1988 年大学毕业后我就一直喜欢贵刊，原因有三：（1）贵刊刊登的文章涉及面广，有宏观方面如植物区系、进化等的内容，也有微观方面的如关于植物结构、植物化学的研究，它们往往都能较好地反映世界和我国植物学主要的研究领域和方向；（2）文章的类型多样，可读性好，如 2001 年发表马金双先生的《云南植物采集史略读后感》、2003 年的《水杉未解之谜的初探》以及近来的《水杉的未尽事宜》，就是很好的例子，读来饶有趣味。我很喜欢《Nature》和《Science》等刊物，原因之一就是读来易懂，融知识性和趣味性于一体；（3）编辑和印刷质量一直很好，我曾自费订阅贵刊并买过增刊。近 20 年来我一直关注着贵刊，认为贵刊编辑、审稿人员是负责的，贵刊为我提供了重要的知识和很多乐趣。

下面就《水杉的未尽事宜》谈点读后感。该文涉及到有关我的两件事：（1）是说我们 2004 年在《北京林业大学学报》（社会科学版）发表了《水杉发现过程的系统研究》一文，但令马先生费解的是我们尽管已见到《云南植物研究》发表的文章《水杉未解之谜的初探》，并在文中加以引用，但仍然把胡先骕 1946 年春见到水杉的事实写成 1946 年秋，这种不负责任的做法不应该发生在我们专门从事水杉研究的学者之中。实际上，胡先骕是 1946 年春还是该年秋见到水杉标本这一问题，我在 1990 年就注意到了。胡先生得知发现水杉活化石的事可能如马先生所说的是 1946 年春或更早，但他是 1946 年 10 月 30 日离开南京到上海，于 11 月 16 日才到达北平的 [见王咨臣等《植物学家胡先骕博士年谱（二）》以及胡先骕先生 1946 年 11 月 17 日的信件（见胡宗刚 2003 年《静生生物调查所与中央林业试验所的两项合作》一文）]。回到北平后，胡先生才可能看到郑万钧先生寄给他的水杉标本。我们上述文中指的是胡先骕先生看到水杉标本的时间。从其行程推测，胡先生看到水杉标本的时间不可能是 1946 年春。（2）是说 2004 年我在《生物多样性》发表《水杉的保护历程和存在的问题》一文时，在正文中引证了马先生的两篇文章，但在文献引证部分均未被列出，这会给读者的阅读和考证带来困难。马先生的这两篇文章十分重要，《生物多样性》编辑部的负责同志和编辑都十分重视，遗憾的是他们后来在最后排印